

111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622

Method and Apparatus For True Priority Based Connection Establishment Within A PNNI ATM Network

Krishna Sundaresan

Daniel Cauchy

Prepared by:

Attorney's Docket No.: 081862.P255

Date of Deposit: November 27, 2001

Janece Shannon

(Typed or printed name of person mailing paper or fee)

(Signature of person mailing paper or fee)

(Date signed)

Method and Apparatus For True Priority Based Connection Establishment Within A PNNI ATM Network

Field of the Invention

[0001] The field of invention relates to networking, generally; and, more specifically, to a method and apparatus for true priority based connection establishment within a PNNI ATM network.

Background

[0002] An exemplary Private Network Node Interface (PNNI) Asynchronous Transfer Mode (ATM) network 101 is shown in **Figure 1**. ATM is a networking technology that transports information with “cells” of data. As such, if a significantly sized body of information (e.g., a document or file) is to be transported across an ATM network, the body of information is effectively “broken down” into a plurality of cells. The plurality of cells are then individually sent across the network and reassembled at the receiving end in order to reconstruct the original body of information.

[0003] The term “connection” or “circuit” is often used to describe a pre-defined path through a network. Typically, when a body of information is to be transported over a network, a connection is setup beforehand that establishes (in some manner and to some extent) the path that the cells will take. Various types of connections may be used within an ATM network 101. These include: 1) permanent virtual circuits (PVCs); 2) switched virtual circuits (SVCs); and 3) soft permanent virtual circuits (SPVCs).

[0004] In the case of PVCs, a quasi-permanent connection is established (e.g., a connection that lasts for days, weeks, months, etc.). PVCs are often used in situations where a large corporate user desires to permanently clear a guaranteed pipe through the network 100 from one large office to another large office. For example, if node 105₁ corresponds to the Customer Premise Equipment (CPE) of a first corporate office and node 105₂ corresponds to the CPE of a second corporate office, a PVC may be established that couples nodes 102₁, 102₄, 102₇ and network lines 103₃, 103₁₁ together (in order to form an end-to-end path through the network 100 between CPEs 105₁ and 105₂).

[0005] Generally, the amount of traffic (e.g., as between two large corporate offices) and the extent of the usage (e.g., every business day for the foreseeable future) justifies the costs associated with dedicating, in a quasi-permanent fashion, a fixed amount of the network's resources to one particular pathway. Typically, a PVC is manually configured by a network manager from a network management control station 104. As such, commands are issued from the network control station 104 to the various nodes in the network 101 that "make up" the PVC (so that the lookup tables, etc. within these nodes can be properly updated).

[0006] Another characteristic of a PVC is that a PVC user simply directs traffic into the network 101 (e.g., from node 105₁) with little or no formal request for transportation services from the network 101. For example, typically, a PVC user at node 105₁ will send ATM cells having the PVC's VPI/VCI across the ATM User Network Interface (UNI) at link 103₁. Based upon the VPI/VCI information, node

102₁ (e.g., as well as subsequent nodes along the PVC path) will be able to properly switch the cells onto a link that corresponds to the PVC path. Thus, because the connection is quasi-permanent and has already been established, there is little or no procedural overhead associated with connection setup (such as a SETUP request message and the like). The user is provided an appropriate VPI/VCI well beforehand (e.g., shortly after PVC setup) which is invoked each time thereafter by the user when the services of the PVC are desired.

[0007] SVCs, on the other hand, are established on a temporary basis rather than a quasi-permanent basis. SVCs efficiently utilize the resources of a network if the network has to support a large number of different connection paths over a fairly brief period of time (e.g., seconds, minutes, hours). In contrast to PVCs, SVCs are usually established on a “call-by-call” basis and therefore have: 1) some form of formal user request to the network 101 for transportation services; and, 2) a connection “setup” procedure that follows the request for transportation services and a connection “tear down” procedure that follows the successful performance of the requested transportation services.

[0008] The connection setup/tear down procedures may be viewed as the “automatic” configuration of a connection within the network rather than manual configuration from a network management control station 104. PNNI is a routing and signaling protocol that determines and establishes connection paths . The PNNI routing protocol is executed on the source endpoint (e.g., source endpoint 102₁ for connections initiated from originating node 105₁), and is often referred to

as a “source” routing protocol. An example of PNNI’s routing and signaling techniques are provided immediately below.

[0009] If node 105₁ (the “originating” node) desires to send information to node 105₂ (the “target” node), the originating node 105₁ will effectively request the network 101 for a connection to be established between nodes 105₁ and node 105₂. Typically, this request takes the form of a SETUP message that is passed over the ATM UNI at link 103₁. The access node 102₁ (which may be referred to as the source endpoint node) receives the SETUP message and determines an appropriate path for the connection through the network via the PNNI routing protocol.

[0010] The SETUP message then traverses the network 101 to the destination endpoint node 102₇. When the SETUP message is received at the destination endpoint node 102₇, a CONNECT message is issued from the destination endpoint node 102₇ to the source endpoint node 102₁. The CONNECT message “bounces”, node-by-node, along the connection path to the source endpoint node 102₁. Each node that receives the CONNECT message updates its lookup table (or other routing/switching platform) with an appropriate reference to the connection being established. When the source endpoint node 102₁ receives the CONNECT message, the VPI/VCI for the connection is passed to the user at the ATM UNI (along link 103₁), the connection is established, and transportation services may commence. After the transportation services are complete, the connection is torndown in a manner similar to that in which it was established.

[0011] An SPVC is often viewed as a blending of an SVC and a PVC. SPVCs are often used to provide guaranteed bandwidth to a particular user (such that the user enjoys service as if a permanent pipe has been established through the network 101) while, simultaneously, the network 101 is allowed to flexibly adapt to different connection paths over brief periods of time (by establishing each SPVC communication with connection setup and tear down procedures). In order to implement an SPVC service, the endpoint nodes of the ATM network 101 (e.g., source node 102₁ and destination node 102₇) are configured to behave like PVC nodes with respect to the user (e.g., along the ATM UNI at link 103₁) while behaving like SVC nodes within the ATM network 101 itself.

[0012] With an SPVC, the source and destination endpoint nodes 102₁ and 102₇ are usually manually configured by the network management station 104 to provide a PVC interface to the users at node 105₁ (and at node 105₂). That is, for example, a quasi permanent VPI/VCI is provided to the user that is to be invoked each time the services of the SPVC are desired. Upon the receipt of ATM cells having this VPI/VCI information, however, the endpoint source node 102₁ triggers the release of a SETUP message which traverses the network 101 to destination endpoint node 102₇. A CONNECT message is returned to the endpoint source node 102₁, and the SPVC is established.

Figures

[0013] The present invention is illustrated by way of example, and not limitation, in the Figures of the accompanying drawings in which.

[0014] **Figure 1** shows an embodiment of a PNNI ATM network.

[0015] **Figure 2** shows an embodiment of a methodology for true priority based connection establishment within a PNNI ATM network.

[0016] **Figure 3** shows an example of a PNNI ATM network's bandwidth resource allocation prior to the establishment of a new connection.

[0017] **Figure 4** shows an example of a PNNI ATM network's bandwidth resource allocation after the new connection referred to with respect to Figure 3 is established.

[0018] **Figure 5** shows an embodiment of a methodology that may be used to help determine a path through an ATM PNNI network at a source endpoint in light of information that pertains to bandwidth resource allocation elsewhere in the network.

[0019] **Figure 6** shows an embodiment of a PNNI Topology State Packet (PTSP).

[0020] **Figure 7** shows an embodiment of a PNNI Topology State Element (PTSE) that may be embedded within the PTSP of **Figure 6**.

[0021] **Figure 8** shows an embodiment of a System Capabilities Information Group (SIG) field that may be embedded within the PTSE of **Figure 6**.

[0022] **Figure 9** shows an embodiment of a node.

Description

[0023] A problem with both SVC and SPVC type connections is that the connection establishment process does not execute a “true” Quality of Service (QoS) priority bumping scheme with respect to the network 101 as a whole. QoS relates to the notion that the various connections supported by the network 101 are to be prioritized in some manner with respect to one another. For example, the network 101 may be configured to give cells associated with higher priority connections lower end-to-end delay across the network 101 than cells associated with lower priority connections.

[0024] Another QoS parameter concerns the connection establishment process. Specifically, under a process that may be referred to as “priority bumping”, higher priority connections are established at the expense of lower priority connections. For example, if the network 101 is supporting a large number of low priority connections at the same moment it happens to receive a request for a high priority connection, one or more of the low priority connections can be “dropped” so that the high priority connection can be established.

[0025] The PNNI routing and signaling scheme is often said to be a “source routing” scheme because the appropriate path for a network connection is determined at the node that acts as its source endpoint (e.g., node 102₁ for connections initiated by node 105₁). Currently, only a limited form of priority bumping is possible. Specifically, only those lower priority connections that actually pass through the source endpoint can be dropped.

[0026] Consequently, a remote node that supports many low priority connections cannot drop these connections in order to support a higher priority connection; unless, these connections also happen to flow through the same endpoint node that seeks to establish the high priority connection. Thus, in a sense, the power to drop one or more lower priority connections in favor of a higher priority connection is localized to the realm of the source endpoint node that seeks to establish the higher priority connection (rather than being distributed over the network 101 as a whole).

[0027] The power to drop connections is limited in the manner described above because the source endpoint node that is responsible for establishing a connection has little or no perspective as to the manner in which bandwidth resources elsewhere in the network have been allocated with respect to connection priority. As such, in order to expand the scope of those lower priority connections that may be dropped in order to support a higher priority connection, a source endpoint should be made aware of the connections being supported on other nodes in terms of their priority and they bandwidth the consume.

[0028] **Figure 2** shows an embodiment of a methodology that accomplishes this task. According to the approach described by the methodology of **Figure 2**, the nodes within an ATM network “broadcast” to one another the bandwidth they have allocated (e.g., in the form of active connections) to each priority level. As such, a source endpoint can develop a full perspective as to the manner in which bandwidth is being allocated, per priority level, and decide in light of this

perspective whether or not the network as a whole can support a newly requested connection.

[0029] Thus, as seen in the methodology of **Figure 2**, a prospective source endpoint node will receive 201 System capabilities Information Group (SIG) information (from the other nodes within the PNNI network in which it resides) that describe how the bandwidth resources (e.g., links, nodes, a portion thereof or a combination thereof) have been allocated with respect to the priority of the connections within the network. SIG information, as described in more detail further below, is a mechanism that has been established by the PNNI scheme for the passing of information that has not been formally provided for by the PNNI standard.

[0030] After the SIG information has been received 201, the prospective source endpoint node will update 202 its present understanding of the network. In various embodiments, the understanding of the network corresponds to the collection of the latest SIG information received from each node within the network. As such, a network wide status is crafted from which those bandwidth resources (e.g., a remote node; and/or, a link or portion of a link that is coupled to a remote node) that can be used to support a connection having a specific priority is identified.

[0031] When a new connection is requested 203 (e.g., formally in the case of an SVC or informally in the case of an SPVC), the prospective source node attempts to determine 204 a possible path through the network. For example, in various "path-recursive" approaches, the prospective source endpoint node is

configured to determine a “first pass” path (e.g., according to a PNNI compliant source routing algorithm) through the network. Then, once the “first pass” path has been determined, the bandwidth resources that would be used along the path are looked into (as understood via the network understanding that has been developed 202) to see if the path can be entertained.

[0032] If the path can be entertained, the connection path is established (which may involve the dropping of lower priority connections) via the issuance 205 of a SETUP message from the source endpoint node. If not, an alternative “second pass” path is determined and analyzed in a similar fashion. The process continues in a recursive fashion until a workable path is identified (in which case the connection is ultimately established); or, alternatively, is not identified (in which case the connection is not ultimately established).

[0033] In other approaches, unlike the approach just described above where different paths are analyzed recursively, a routing algorithm may be used that automatically determines the appropriate path. For example, a weighted routing algorithm may be used that assigns weights to the bandwidth resources of the network that are available for the priority level of the requested connection. The mathematics of the algorithm then automatically finds the path having the heaviest (or lightest) weight as it progresses systematically from the source endpoint node to the destination endpoint node. Other routing approaches may be possible as well.

[0034] Figures 3 and 4 demonstrate an embodiment of the organization and/or presentation of the SIG information that is broadcast around the network

101 of **Figure 1**. As the SIG information is broadcast within the network 101, it will eventually be collected by a prospective source endpoint such as node 102₁. The discussion pertaining to Figures 3 and 4 relate to a path recursive source routing technique. However, it is important to point out that the organization and presentation of the SIG information as depicted in **Figures 3 and 4** can be used for other routing techniques as well (e.g., such as the weighted technique alluded to just above).

[0035] **Figure 3** shows a portion of an exemplary understanding that has been developed for network 101 of **Figure 1** by prospective source endpoint 102₁. As described above, an understanding of a network may be crafted by collecting the latest SIG information units that have been broadcast within the network for various bandwidth resources. A bandwidth resource is a network element that transports traffic such as a node or link or portion thereof or combination thereof. The SIG information units that are shown in **Figures 3 and 4** depict bandwidth allocations, on a per priority level basis, for some of the links within network 101 of **Figure 1**.

[0036] Specifically, **Figure 3** shows exemplary SIG information units that have been broadcast for links 103₂, 103₃, 103₅, 103₁₁, 103₁₂, and 103₁₄ of **Figure 1**. As a brief aside, note that only link bandwidth information is observed in **Figure 3** rather than nodal bandwidth. Generally, the particular bandwidth resources that are described by the SIG information units will depend upon the needs of the particular routing technique. That is, various routing techniques may use only link related bandwidth information whereas various other routing algorithm

techniques may use only nodal related bandwidth information. Further still, alternate routing algorithm techniques may use some combination of nodal and link related bandwidth information.

[0037] With regard to the discussion of **Figures 3 and 4**, however, one may assume that the applicable routing technique uses only link bandwidth information. In the exemplary embodiment of **Figures 3 and 4**, there are “n” priority levels. Thus, as seen in **Figures 3 and 4**, each SIG information unit includes a “total” bandwidth description (that describes the total bandwidth of the link being described), as well as n priority level bandwidth descriptions (that each describe the bandwidth allocated to the corresponding priority level).

[0038] Referring to **Figures 1 and 3**, consider an example in which user node 105₁ requests endpoint node 102₁ to establish a 155.5 Mb/s, level 3 priority SVC or SPVC connection to user node 105₂ (through endpoint node 102₇). As such, endpoint node 102₁ acts as the source endpoint node for the requested connection. Assuming that the path-recursive routing protocol being executed on the source endpoint node 102₁ identifies the 102₁-103₃-102₄-103₁₁-102₇ path through the network 101 as the “first pass” path, note that the SIG information for links 103₃ and 103₁₁ illustrate that link 103₃ can entertain the requested connection but that link 103₁₁ cannot.

[0039] That is, the SIG information unit for link 103₃ indicates that 311 Mb/s of that link’s bandwidth has been allocated for connections having a priority level of 1 or a priority level of 2 (i.e., 155.5 Mb/s of priority level 1 traffic + 155.5 Mb/s of priority level 2 traffic = 311 Mb/s of priority level 1 and 2 traffic). Because the

requested connection has a priority level of 3 (and because higher priority connections have lower priority level numbers), the 311 Mb/s worth of allocated bandwidth will not be dropped from link 103₃ (because this bandwidth has been dedicated to higher priority connections) and may be said to be "unavailable".

[0040] Nevertheless, 155.5 Mb/s of the link's bandwidth is freely available (i.e., 466.5 Mb/s of total bandwidth - 311 Mb/s of unavailable bandwidth = 155.5 Mb/s of available bandwidth). As such, link 103₃ can be used to support the requested connection. Link 103₁₁, however, has no bandwidth available for links of priority level 3 or lower. To first order, all 466.5 Mb/s of the link's bandwidth has been consumed; and, none of the priority levels for which bandwidth has been allocated is lower than the priority level of the requested connection. As such, there are no lower priority connections that can be "bumped" in order to make room for the requested connection. Link 103₃ is therefore deemed unavailable for the connection and the "first pass" path through the network 101 can not be established.

[0041] If the routing protocol identifies a 103₂-102₂-103₅-102₅-103₁₂-102₇ path through the network as the "second pass" path, this second pass path can be entertained. Specifically, each of links 103₂, 103₅, and 103₁₂ are only carrying traffic having a lower priority (priority levels 4, 5 and 6) than the requested connection (priority). As such, in an embodiment, any or all of the connections on these links may be dropped to support connections having a priority level of 3 or higher.

[0042] In a further embodiment, connections are dropped according to reverse priority. That is, the lowest priority connection(s) beneath the priority level of the requested connection are dropped first, followed by the next lowest, etc. **Figure 4** shows a depiction of the broadcast SIG information after the requested connection is established on the 103₂-102₂-103₅-102₅-103₁₂-102₇ network path. Note that, consistent with the reverse priority bumping scheme, the priority level 3 connection that was just requested has been established at the expense of the priority level 6 connection(s) (that were shown in **Figure 3**).

[0043] That is, the 155.5 Mb/s worth of bandwidth that was originally allocated to the priority level 6 connection(s) has been re-allocated to the newly established priority level 3 connection. As a result, the priority level 4 and priority level 5 connection(s) have been unaffected by the establishment of the new, priority level 3 connection. Note that the newly established priority level 3 connection has also dropped all of the previous traffic that existed on access link 103₁₄.

[0044] **Figure 5** shows an embodiment of the methodology described above that can be used to determine if a particular bandwidth resource can support a requested connection. Firstly, any freely available bandwidth is added 501 to the bandwidth that has been allocated to lower priority connection(s) in order to determine the amount of bandwidth that may be consumed by a newly requested connection. If this bandwidth available to the newly requested connection is not greater than or equal to the bandwidth to be consumed by the requested

090645 134
T022T 5349550

connection the connection cannot be entertained 503 at the bandwidth resource being analyzed.

[0045] If the bandwidth available for the newly requested connection is greater than the bandwidth to be consumed by the newly requested connection, the requested connection may be entertained at the bandwidth resource being analyzed. Furthermore, if the freely available bandwidth is greater than or equal to the bandwidth to be consumed by the requested connection, then the connection may be established 506 without any priority bumping.

[0046] If, however, the freely available bandwidth is not greater than or equal to the bandwidth to be consumed by the requested connection, the connection is established 506 with priority bumping (e.g., at the expense of the lowest priority connection(s)). When a connection is being established, the source endpoint node issues a SETUP message that traverses the network to the destination endpoint node. In an embodiment, the SETUP message is configured to include the priority level and the bandwidth of the connection being established so that the nodes that carry the new connection can determine which connections are to be dropped.

[0047] Referring back to **Figure 2**, recall that SIG information is broadcast by each node in the network so that a prospective source endpoint node can receive 501 the SIG information and develop an understanding 502 of the network's bandwidth allocations. As part of the PNNI scheme, each node within the network is typically designed to "broadcast" other information (i.e., other than SIG information) that pertains to its understanding of itself and/or the network in which

it resides. These broadcasts may occur at specific time intervals and/or upon the occurrence of certain special events.

[0048] For example, referring to **Figure 1**, if a node 102₅ observes that networking link 103₁₀ is not working, the node 102₅ will broadcast this event to its neighboring nodes 102₂, 102₇. Upon the reception of this information, the neighboring nodes 102₂, 102₇ will "update" their internal understandings of the network (to reflect this event) as well as rebroadcast this event to their neighboring nodes so that they may update their internal understandings as well. The information is continually rebroadcast as appropriate so that the affected nodes can update their understandings of the network and behave accordingly.

[0049] Thus, in a sense, the occurrence of the event ripples through the network so that its constituent nodes can cohesively route information around the downed link 103₁₀ in response. In other cases, typically, the network's nodes 102₁ through 102₇ are also configured to broadcast current status information as well as special events. Thus, on a broader scale, the nodes of the network may be said to communicate procedural (e.g., "control") information with one another as well as the substantive information associated with user traffic.

[0050] This control information is often organized into one or more PNNI Topology State Elements (hereinafter, referred to as PTSEs) that are embedded into a PNNI Topology State Packet (hereinafter, referred to as a PTSP). A PTSP is a packet that acts as the broadcast mechanism while a PTSE acts as a component of the PTSP's payload. Thus, for example, if a node has information to broadcast it issues a PTSP that carries one or more PTSEs that each have

the information to be communicated. An embodiment 600 of a PTSP is shown in **Figure 6** and an embodiment 701 of a PTSE is shown in **Figure 7**.

[0051] Referring to **Figure 6**, a PTSP may be viewed as having a header field 606 and a PTSE field 601. The header field 606 has various header information (e.g., checksum info, lifetime, etc.) as well as the identification of the node that is issuing the PTSP (which is located within the originating node ID field 603), the peer group within which the originating node resides (which is located within the Peer Group ID field 604). PNNI Peer groups are discussed in more detail toward the end of this description.

[0052] The PTSE field 601 includes one or more PTSEs 601₁ through 601_x. An embodiment 701 of a PTSE is shown in Figure 7. That is, for example, the PTSE embodiment 701 of **Figure 7** may be viewed as corresponding to the PTSE 601₁ of **Figure 6**. Referring to **Figure 7**, note that a PTSE may also be viewed as having a header field 702 and a payload field 603. The header field 702 includes various header information such as a type field 706 that identifies the data structure 701 as a PTSE, a length field 707 that identifies the length of the PTSE, a reserved field 709 for potential future uses and a checksum field 712.

[0053] The PTSE header field 702 also includes a identifier field 710 that identifies the type of PSTE that PTSE 701 corresponds to. That is, PNNI employs a characterization scheme so that specific types of information can be binned together or recognized within a common PTSE format. The various PTSE types include (among possible others): 1) Horizontal Link; 2) Uplink; 3)

External Address; 4) Internal Address; 5) Nodal Parameters (complex node); and 6) Nodal. Those of ordinary skill can identify the purpose and/or use of each PTSE type.

[0054] Referring to the PTSE embodiment 701 of **Figure 7**, note that the payload field 703 may be viewed as being partitioned into an “industry standard” field 704 and the aforementioned System Capabilities Information Group (SIG) field 705. The industry standard field 704 is used to carry specific information according to a specific format that has been articulated by the PNNI standard. The SIG field 305, by contrast, is used for developers of PNNI compliant networking gear that seek to include special features beyond those recognized or articulated by the PNNI standard.

[0055] Through the use of the SIG field 705, two nodes from the same manufacturer can communicate information with one other that is not specifically provided for by the PNNI standard; while, at the same time, operate in compliance with the PNNI standard. That is, those nodes that can understand and use the contents of the SIG field 705 may do so while those that do not understand the SIG field 705 contents may simply ignore its information (as well as forward the PTSE having the SIG field to another node via a rebroadcast effort).

[0056] The Horizontal Link PTSE type is commonly used to transport information that pertains to a link, or a portion of a link. That is, finer granularities that the whole of a link’s resources may be specified or described with a Horizontal Link PTSE. These finer granularities may be used to develop a

deeper understanding of the network's bandwidth allocations. For example, the industry standard field 704 of a Horizontal Link PTSE can specify a particular type of service such as a Constant Bit Rate (CBR) service, a real time or non-real time Variable Bit Rate (VBR) service, an Available Bit Rate (ABR) service and an Unspecified Bit Rate (UBR) service.

[0057] As such, the priority levels of a particular link may be further specified so as to define specific priority levels for each of the service types that exist on the link. As such, the network can be configured to bump lower priority connections of a specific type in response to a request for a connection of the same service type. Further still, the industry standard field 704 of a Horizontal Link PTSE can specify particular QoS parameters (e.g., average cell rate, max cell rate, cell transfer delay, cell delay variation, and cell loss ratio).

[0058] As such, the priority levels of a particular link may be further specified so as to define specific priority levels not only for each service type, but also for each common combination of QoS parameters that exist on the link. Thus, in various embodiments, the industry standard field 704 of a Horizontal Link PTSE is used to specify the service type or the service type and specific QoS settings for which the bandwidth allocations are detailed according to different priority levels. In various embodiments, the bandwidth allocations for the various priority levels of each of these finer granularities of a link may be specified within the SIG field 705.

[0059] **Figure 8** shows an embodiment 805 of a SIG field. That is, the SIG field 805 of **Figure 8** may be viewed as an embodiment of the SIG field 705 of

Figure 7 that can be used to express the address change of an SPVC endpoint.

The SIG field embodiment 805 of **Figure 8** can also be viewed as having a header field component 801 and a payload field component 802.

[0060] The header field component 801 includes various header information such as a type field 806 (that indicates the data structure 805 is a SIG field), a length field 807 that describes its length and an Organization Unique Identifier (OUI) field 808 that is typically used to recognize the manufacturer of the node that issued the SIG information (i.e., is a “vendor-specific” label). As a SIG field is typically used by the nodes of a common manufacturer to support functional improvements (beyond the PNNI standard) that are unique to the products of the manufacturer, the OUI field 808 is often used by a node to decide whether or not to ignore a received SIG field. That is, if the vendor specific label of the OUI field 808 “matches” the vendor of the node that receives the SIG information, the SIG information will be “looked into”; otherwise, the SIG information will be discarded.

[0061] Within the payload 802 of the SIG field 805, the ID # field 803 identifies the particular type of information being delivered by the SIG 805. This allows a node that supports vendor-specific functionality to understand the specific type of information enclosed in the payload 802. As such, in an embodiment, a specific binary number is used to identify that the SIG field 805 includes information related to the priority levels of the particular service type (and perhaps QoS parameter combination) that are specified in the industry standard portion of the PTSE that carries the SIG field 805. In the particular embodiment of **Figure 8**,

the bandwidth allocations made to each priority level are specified in the priority reference and bandwidth allocation fields 804₁ through 804_n.

[0062] The rate at which PTSE information (having priority level bandwidth distributions embedded within its SIG field) is broadcast from a particular node may vary from embodiment to embodiment. For example, in one embodiment, PTSE information may be broadcast for each new connection that reflects a change in the distribution of bandwidth resources to the various priority levels. In various networking environments, however, issuing new PTSE information from each node that undergoes a mere change in distribution may congest the network with PTSP packets; or, may make an already congested network even further congested.

[0063] As a network gets congested, the delay experienced by its cells increases. Thus, the network understandings that are developed by the network's endpoint nodes may become in danger of being unacceptably inaccurate. That is, the SIG information embedded within the most recently received PTSP packets (which have experienced significant delay) is stale with respect to the actual state of the bandwidth resources they are describing.

[0064] As such, in alternate embodiments, the conditions that trigger the release of PTSE information having priority related SIG information may be scaled back. For example, rather than release PTSE information with each change in the distribution of bandwidth resources, in an alternate embodiment PTSE information is released with each "significant" change in the distribution of bandwidth resources. As just one approach, if any priority level experiences a

looked-for percentage change (within a finite time period) of allocated bandwidth in either direction (e.g., a 50% increase in bandwidth resources or a 50% decrease in bandwidth resources), PTSE information reflecting the change is broadcast.

[0065] In this case, the degree of “significance” is defined by the looked for percentage. As such, other embodiments can configure other degrees of significance such as 25%, 33%, or 66% to name just a few. Note that the degree of significance can adapt to network conditions. For example, under light congestion conditions, the network can reliably handle more PTSP packets. As such, the degree of significance may be lowered (e.g., from X% to Y% where $X > Y$) so that more PTSE information is issued per unit time (which increases the accuracy of the network understandings made at the endpoint nodes). As network congestion increases, the degree of significance may be increased (e.g., from Y% to X%, where $X > Y$) so that the PTSE information does not significantly add to the network’s congestion.

[0066] In other embodiments, rather than trigger the release of PTSE information having priority related SIG information only on an event (such as a change in bandwidth allocation or a significant change in bandwidth allocation), it may alternatively (or in combination) be released periodically. Again, the rate at which the PTSE information is released may vary with network congestion levels. For example, by issuing PTSE information less frequently as congestion rises; or, more frequently as congestion falls.

[0067] Note that, to the extent that the understandings of the network that are being maintained by the endpoint nodes become inaccurate, the “crankback” mechanism associated with PNNI signaling may be employed to recover from such an inaccuracy. Specifically, an inaccurate network understanding may result in the release of a SETUP message from a source endpoint node for a connection path that cannot be entertained because higher priority connections have already been established between the time the source endpoint node’s latest PTSE information was issued and the time the connection request was received.

[0068] Upon the receipt of such a SETUP message by a node that is intended to carry the new connection yet cannot support it (because its bandwidth resources are already consumed by higher or equal priority level connections), the node may return a “crankback” message back to the source endpoint node that issued the SETUP message. The crankback message can be configured to contain information that effectively explains the problem to the source endpoint node. In response, the source endpoint node can update its network understanding and re-determine another path through the network.

[0069] As routing and signaling protocols are often implemented with software, it is to be understood that embodiments of this invention may be used as or to support a software program executed upon some form of processing core (such as the CPU of a computer) or otherwise implemented or realized upon or within a machine readable medium. A machine readable medium includes any mechanism for storing or transmitting information in a form readable by a

machine (e.g., a computer). For example, a machine readable medium includes read only memory (ROM); random access memory (RAM); magnetic disk storage media; optical storage media; flash memory devices; electrical, optical, acoustical or other form of propagated signals (e.g., carrier waves, infrared signals, digital signals, etc.); etc.

[0070] Furthermore, it is noteworthy to point out that a network node (which may also be referred to as a networking node, a node, a networking system and the like) is a system designed to act as a switch or a router or other device that relays information from a first networking line to a second networking line. A depiction of a networking node 900 is observed in **Figure 9**. A plurality of networking lines 901₁ through 901₆ (e.g., copper cables or fiber optic cables) are shown in **Figure 9** as being coupled to the networking node 900.

[0071] The node 900 is mostly responsible for collecting a traffic unit (e.g., a packet, a cell or a Time Division Multiplexed (TDM) time slot) from a first networking line (e.g., networking line 901₁) and re-transmitting at least a portion of it (e.g., its payload and various sections of its header) onto a second networking line (e.g., networking line 901₆). As such, the node 900 effectively relays information so that it may be carried over various geographic distances. Some degree of intelligence is involved in the relaying process so that the traffic units being collected are forwarded onto an appropriate networking line (e.g., in light of their source address and destination address).

[0072] As such, the node 900 of **Figure 9** shows an traffic ingress/egress layer 902 and a switching/routing layer 903. The ingress/egress layer 902 is responsible for collecting inbound traffic units from the networking lines upon which they arrived; and, presenting at least a portion of them (e.g., their header information) to the switching/routing layer 903. The ingress/egress layer 902 is also responsible for transmitting outgoing traffic units onto a networking line in response to the direction or control of the switching/routing layer 903.

[0073] The switching/routing layer 903 is responsible for effectively deciding which networking line is an appropriate networking line upon which a particular traffic unit should be transmitted upon. The switching/routing layer 903 often performs this activity based upon header information or other control information (such as SS7 based TDM connection information) associated with each traffic unit. Connection establishment and tear-down procedures (as well as network topology broadcasts or other networking overhead information) can often be viewed as being integrated into (or coupled to so as to communicate with) the switching/routing layer 903.

[0074] Note that the architecture of a networking system having a routing/switching layer 903 and an ingress/egress layer 902 may vary from embodiment to embodiment. For example, in some cases the switching/routing layer 903 may be designed onto a single card; or, in other cases, the switching/routing layer 903 may be designed across a plurality of cards. Also, in some cases the switching/routing layer 903 (or a portion thereof) may be

integrated onto a Line Interface Card (LIC) that also acts as part of the ingress/egress layer 902.

[0075] In the foregoing specification, the invention has been described with reference to specific exemplary embodiments thereof. It will, however, be evident that various modifications and changes may be made thereto without departing from the broader spirit and scope of the invention as set forth in the appended claims. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense.

090943 12701
FOZET 5349550